Center
for
Economic Research

No. 2000-120

DOUBLE CHECKING FOR TWO ERROR TYPES

By V.M. Raats and J.J.A. Moors

December 2000

ISSN 0924-7815

Double checking for two error types

V.M. Raats[¤] and J.J.A. Moors*

December 11, 2000

Abstract

Auditing a large population of recorded values is usually done by means of sampling. Based on the number of incorrect records that is detected in the sample, a point estimate and a con…dence limit for the population fraction of incorrect values can be determined. In general it is (implicitly) assumed that the auditor does not make mistakes while judging the correctness of the values. However, in practice this assumption does not necessarily hold: auditors are human and can make errors. To take this possibility into account, a subsample of the audited records is checked once more by a second auditor who is assumed never to make mistakes. The information obtained from these two samples should be combined to derive an estimate for the error rate in the population.

The starting point for this type of double checking was Moors et al. (2000). Only one possible error type was considered: auditors could only miss (fail to detect) existing errors. For the case of random sampling, the maximum likelihood estimator as well as an upper con…dence limit for the error rate were derived.

The present paper gives extensions in two directions. Firstly, a second error type is introduced: the auditor may consider a correct value as an error. Again, the sample information of both auditor and infallible expert is combined to give point and interval estimates for the fraction of errors in the population. Secondly, a Bayesian analysis is presented for both the model with one error type and the extended model.

Key words: auditing, Bayesian analysis, con…dence limit, double inspection, error types, inspecton errors, quality control

Jel codes: C11, C13, C42, C63, M41

[¤] Tilburg University, P.O. Box 90153, 5000 LE Tilburg

1 Introduction

In a random sample of 500 social security payments, sixteen errors were found by an auditor. Since Dutch social rules and regulations are notoriously complicated, the possibility could not be excluded that the auditor's work still contained errors. Hence, a subsample of 53 of the checked records was checked once more - now by an expert assumed to be infallible. The expert found one additional error. This problem was analyzed in MOORS et al. (2000). Since their model will be extended here, it will be summarized now.

Consider a large population in which a fraction p_1 of the values is incorrect. In a random sample of n records the auditor judges X elements to be incorrect (random variables are denoted in capitals, their realizations in lower case). However the possibility exists that the auditor misses an error: with a probability p_2 an incorrect element is (erroneously) judged to be correct. Hence, a second auditor who is assumed to be faultless (the expert) checks a subsample of the records, of size m, once more. In this subsample the expert now determines the real number of incorrect values (Y); the number of these errors missed by the auditor is denoted by Y_1. Further, the total number X of errors found by the auditor can be split up in a number X_1 from the double checked subsample, and X_2 among the $n - m$ remaining records ($X = X_1 + X_2$). Table 1 shows the information obtained from both checks.

Table 1. Repeated audit control with one error type

Double checked sample				Single checked sample	
	Expert				
First auditor	Correct	Incorrect	Total	First auditor	
Correct	$m - Y$	Y_1	$m - X_1$	Correct	$n - m - X_2$
Incorrect	0	X_1	X_1	Incorrect	X_2
Total	$m - Y$	Y	m	Total	$n - m$

Note that $(Y_1, X_1, m - Y)$ has the multinomial distribution $M(m; p_1p_2; p_1(1 - p_2); 1 - p_1)$. The full information from both samples is contained in the triplet (Y, Y_1, X_2); its joint probability distribution reads

$$
\begin{cases}
L(Y) = B(m; p_1) \\
L(Y_1 | Y = y) = B(y; p_2) \\
L(X_2) = B(n - m; p_1(1 - p_2)) \\
X_2 \text{ and } (Y; Y_1) \text{ independent.}
\end{cases}
\tag{1}
$$

where L indicates the distributional law of a random variable. Note that $p_3 = p_1(1 - p_2)$ denotes the probability that the auditor finds a record to be in error; hence $L(X) = B(n, p_3)$.

In MOORS et al. (2000) the maximum likelihood (ML) estimator

$$G_1 = \frac{(m - Y)X + nY_1}{n(m - X_1)}$$

for p_1 was derived. For $X_1 = m$ however, this estimator breaks down. In this case the loglikelihood equation reduces to

$$\log L(p_1, p_3) = c + (n - x)\log(1 - p_3) + x \log p_3$$

so that only the ML estimator $G_3 = X/n$ is unique. Rather arbitrarily, the value $G_1 = 1$ was chosen in this special case, leading to the revised ML estimator

$$G_1 = \begin{cases} \dfrac{(m - Y)X + nY_1}{n(m - X_1)} & \text{for } 0 \leq X_1 < m \\ 1 & \text{for } X_1 = m \end{cases} \tag{2}$$

Note that for all relevant p_1-values, $\Pr(X_1 = m)$ was negligible, so that the practical consequences of this choice were nil.

Further, following COX & HINKLEY, it was argued that an upper (1-α)-confidence limit p_1^u for a given realization g_1 is obtained from

$$p_1^u = \max_{p_1} \{p_1, p_2 : \Pr(G_1 \leq g_1 | p_1, p_2) \geq \alpha\} \tag{3}$$

A MATLAB-program was developed to calculate p_1^u; an example will be presented in Section 2.3.

The organization of the paper is as follows. Section 2 describes the extended model which not only considers the possibility of overlooking errors (with probability p_2), but also the possibility of making up errors: with probability p_4 the auditor judges a correct value to be in error. Maximum likelihood estimators for these three parameters are derived and a method to obtain an upper limit for p_1 is presented. The estimates and upper limit are calculated for an example which shows that especially the possibility of missing errors causes the upper limit for p_1 to increase sizeably, while the impact of making up errors on the upper limit is considerably less. Section 3 and 4 discuss the Bayesian approach of both models. This approach leads in general to lower upper limits for p_1 than the classical approach, particularly because in the Bayesian approach a weighted average of all different values of p_2 (and p_4) is taken into consideration, and not only the worst case as in the classical approach. The final Section 5 discusses the main results and gives some conclusions and possible extensions.

2 Two error types

2.1 The model

The model discussed in the Introduction has the disadvantage of being asymmetrical: auditors may judge erroneous records to be correct, but not the other way around. A logical step is to extend the previous model by allowing the auditor to make up errors, with probability p_4. So, now the 'quality of an auditor' is expressed by means of two (conditional) probabilities:

$$p_2 = \Pr\{\text{error is seen as correct}\}$$
$$p_4 = \Pr\{\text{correct value is seen as erroneous}\}$$

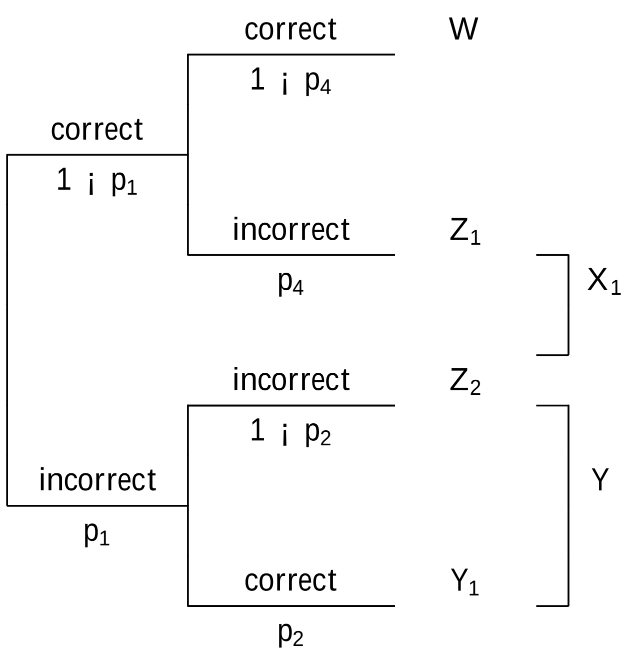
Again, a subsample of the records checked by the auditor, is evaluated once more by an infallible expert. Now, the latter may also find records viewed as erroneous by the auditor to be correct after all. The notation of the previous section is extended accordingly: in the double checked subsample, the number X_1 of errors found by the auditor is split up in Z_2 true errors and Z_1 made up errors. If W denotes the number of correct records, correctly evaluated by the auditor, Table 1 should be replaced by Table 2. Figure 1 presents the double checked subsample in an alternative way.

Table 2. Repeated audit control with two error types

Double checked sample				Single checked sample	
	Expert				
First auditor	Correct	Incorrect	Total	First auditor	
Correct	W	Y_1	$m - X_1$	Correct	$n - m - X_2$
Incorrect	Z_1	Z_2	X_1	Incorrect	X_2
Total	$m - Y$	Y	m	Total	$n - m$

Figure 1. Repeated audit control with two error types

Population First auditor Number



The foursome $(Y_1, Z_2, Z_1; W)$ has the multinomial distribution $M(m; p_1p_2; p_1(1-p_2); (1-p_1)p_4; (1-p_1)(1-p_4))$. The full information of the double checked subsample can be expressed by the triplet (Y, Y_1, Z_1). Combining this with X_2, the joint probability distribution of all sample results is determined by the following distributions:

$$\begin{cases} L(Y) = B(m; p_1) \\ L(Y_1 | Y = y) = B(y; p_2) \\ L(Z_1 | Y = y) = B(m - y; p_4) \\ L(X_2) = B(n - m; p_1(1-p_2) + (1-p_1)p_4) \\ X_2 \text{ independent of } (Y; Y_1; Z_1) \end{cases} \tag{4}$$

Introducing $p_5 = (1-p_1)p_4$ for the probability that the auditor observes a correct record and judges it to be in error, the distribution of X, for example, may be written as $L(X) = B(n; p_3 + p_5)$.

2.2 Point estimators

Note that four distributions are needed to describe the complete sample information on the three unknown parameters. Hence, there are several ways to obtain moment estimators. The most plausible moment estimators for p_1, p_2 and p_4 follow from the expectations

$$E(Y_1) = m p_1 p_2;\; E(Z_1) = m(1 - p_1)p_4;\; E(X) = n(p_3 + p_5)$$

The resulting moment estimator F_1 for p_1 reads

$$F_1 = \frac{X}{n} + \frac{Y_1 - Z_1}{m}$$

It has the property that the numbers of the two different errors compensate each other: if $Y_1 = Z_1$, the estimator reduces to the usual sample fraction of errors. Besides, F_1 is not constrained to the interval $[0,1]$. Both properties are not very satisfactory.

To find the maximum likelihood (ML) estimator, the loglikelihood function is derived from (4); in terms of (p_1, p_3, p_5) it reads:

$$\begin{aligned} \log L(p_1, p_3, p_5) = \;& c + y_1 \log(p_1 - p_3) + z_2 \log p_3 + z_1 \log p_5 + w \log(1 - p_1 - p_5) \\ & + x_2 \log(p_3 + p_5) + (n - m - x_2) \log(1 - p_3 - p_5) \end{aligned}$$

It will be assumed first that w, y_1, z_1 and z_2 are all positive. Equating the three partial derivatives to 0 leads to three equations for the ML estimates g_i for p_i $(i = 1, 3, 5)$:

$$\left.\begin{array}{ll} \text{(a)} & \dfrac{y_1}{g_1 - g_3} = \dfrac{w}{1 - g_1 - g_5} \\[2ex] \text{(b)} & \dfrac{y_1}{g_1 - g_3} - \dfrac{z_2}{g_3} = \dfrac{x_2}{g_3 + g_5} - \dfrac{n - m - x_2}{1 - g_3 - g_5} \\[2ex] \text{(c)} & \dfrac{w}{1 - g_1 - g_5} - \dfrac{z_1}{g_5} = \dfrac{x_2}{g_3 + g_5} - \dfrac{n - m - x_2}{1 - g_3 - g_5} \end{array}\right\}, \tag{5}$$

This system can be solved as follows. First of all, (5a)-(5b)+(5c) reduces to

$$z_2\, g_5 = z_1\, g_3 \tag{6}$$

while (5a) is equivalent to

$$y_1(1 - g_3 - g_5) = (w + y_1)(g_1 - g_3) \tag{7}$$

Substitution of (6) and (7) into the right-hand side of (5b) yields after some simplification

$$x_1 y_1 (n - x) g_3 = x(w + y_1) z_2 (g_1 - g_3) \tag{8}$$

Using (6), (5a) can be rewritten as

$$y_1(z_2 - x_1 g_3) = (w + y_1)z_2(g_1 - g_3) \tag{9}$$

Finally, combination of (8) and (9) gives

$$g_3 = \frac{x\, z_2}{n\, x_1}$$

This expression even holds for $y_1 = 0$; the only exception is of course the case $x_1 = 0$: Ignoring this exception for the moment, the ML estimators for the auxiliary variables become

$$G_3 = \frac{XZ_2}{nX_1}; \quad G_5 = \frac{XZ_1}{nX_1} \tag{10}$$

In principle, the main estimators follow immediately:

$$\begin{cases} G_1 = \dfrac{nX_1Y_1 + X(WZ_2 - Y_1Z_1)}{nX_1(m - X_1)} \\[2ex] G_2 = \dfrac{(n - X)X_1Y_1}{nX_1Y_1 + X(WZ_2 - Y_1Z_1)} \\[2ex] G_4 = \dfrac{X(W + Y_1)Z_1}{nX_1W - X(WZ_2 - Y_1Z_1)} \end{cases} \tag{11}$$

Note that for $Z_1 = 0$, the formulae for G_1 and G_2 reduce to expression (6) in Moors et al. (2000).

Like in Section 1, this derivation breaks down into several cases. For $Y = 0$, the subsample contains no incorrect values, so that no information on p_2 is obtained. Similarly, for $Y = m$, no information on p_4 is available.

In the cases $X_1 = 0$ and $X_1 = m$ the ML estimator for G_1 breaks down. As before the loglikelihood does not lead to a unique ML estimator for p_1; again, somewhat arbitrary values were chosen. Details can be found in MOORS(1999).

In terms of $(X; X_1; Y; Y_1)$; our final ML estimator for p_1 is given by

$$G_1 = \begin{cases} \dfrac{Y_1}{m} & \text{for } X_1 = 0 \\[2ex] \dfrac{(n - X)Y_1}{n(m - X_1)} + \dfrac{X(Y - Y_1)}{nX_1} & \text{for } 0 < X_1 < m \\[2ex] \dfrac{Y - Y_1}{m} & \text{for } X_1 = m \end{cases} \tag{12}$$

A more intuitive understanding of (11) is obtained by considering the two inverse conditional probabilities

$$\begin{aligned}&\Pr(\text{incorrect record}\,|\,\text{auditor judges record to be correct})\\&\quad = p_1p_2/[p_1p_2 + (1-p_1)(1-p_4)];\end{aligned}$$

$$\begin{aligned}&\Pr(\text{incorrect record}\,|\,\text{auditor judges record to be incorrect})\\&\quad = p_1(1-p_2)/[1-p_1p_2-(1-p_1)(1-p_4)]\end{aligned}$$

Substituting the estimators in (11) for the p_i gives for these inverse probabilities the logical estimators

$$Y_1/(m-X_1);\ Z_2/X_1$$

(The former holds as well in the one possible error situation.) Now, a simple interpretation of G_1 follows by rewriting (11) as

$$G_1 = [Y + (n-m-X_2)\frac{Y_1}{m-X_1} + X_2\frac{Z_2}{X_1}]/n$$

For example, the last term between the brackets is the estimated number of true errors among the records that were considered erroneous by the auditor.

Appendix A shows that the distribution of the ML-estimator for the fraction of errors in the population (12) is symmetrical with respect to the point $(p_2, p_4) = (0.5, 0.5)$. (The intuitive explanation is that for high values of p_2 and p_4, all the auditor's judgements should be reversed: 'correct' is better interpreted as 'incorrect', and vice versa.)

The (1-α)-upper limit p_1^u for a given value of the ML-estimator can be found by adding the additional nuisance parameter p_4 to (3):

$$p_1^u = \max_{p_1}\{p_1, p_2, p_4 : \Pr(G_1 \leq g_1|p_1, p_2, p_4) \geq \alpha\} \tag{13}$$

The calculation in practice is illustrated in the next section.

2.3 Examples

Since there were no real life data available which contained made up errors, the point estimates and upper limit are calculated for some fictitious numerical data. These data are closely related to the practical data which were used in MOORS et al. (2000). The assumption that the double checked sample contains one extra made up error leads in combination with the practical data to the following numerical example which RAATS (1999) examined.

Table 3. Numerical example

Double checked sample			
	Expert		
First auditor	Correct	Incorrect	Total
Correct	$w = 49$	$y_1 = 1$	50
Incorrect	$z_1 = 1$	$z_2 = 2$	$x_1 = 3$
Total	50	$y = 3$	$m = 53$

Single checked sample	
First auditor	
Correct	433
Incorrect	$x_2 = 14$
Total	n ¡ $m = 447$

For this example, (11) results in the ML estimates

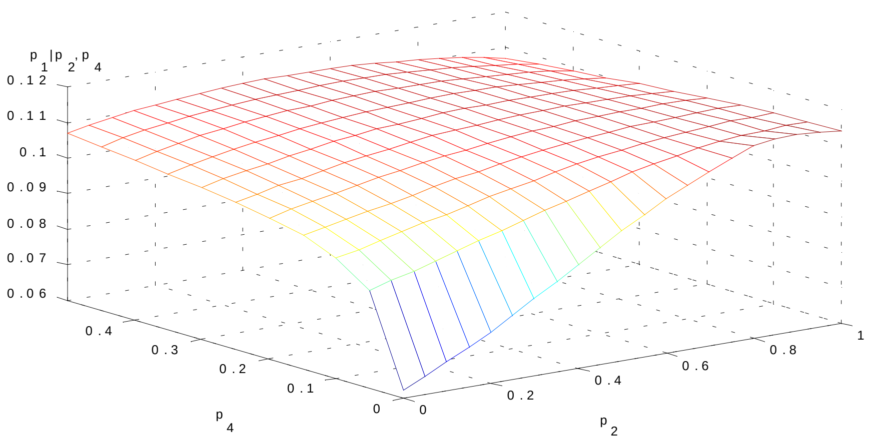
$$g_1 = 0{:}042, \quad g_2 = 0{:}460; \quad g_4 = 0{:}012$$

To determine the accompanying upper 95%-con…dence limit p_1^u, the quantity

$$p_1^u j p_2; p_4 = \max_{p_1} f p_1 : \Pr(G_1 \cdot 0{:}042 j p_1; p_2; p_4) \text{ ¸ } 0{:}05 g$$

has to be calculated for all possible values of p_2 and p_4. Thanks to the symmetry of G_1 with respect to the point $(p_2; p_4) = (0{:}5; 0{:}5)$, the calculations may be limited to the p_4 interval [0, 0.5]. Figure 2 gives a 3-dimensional illustration of the results.

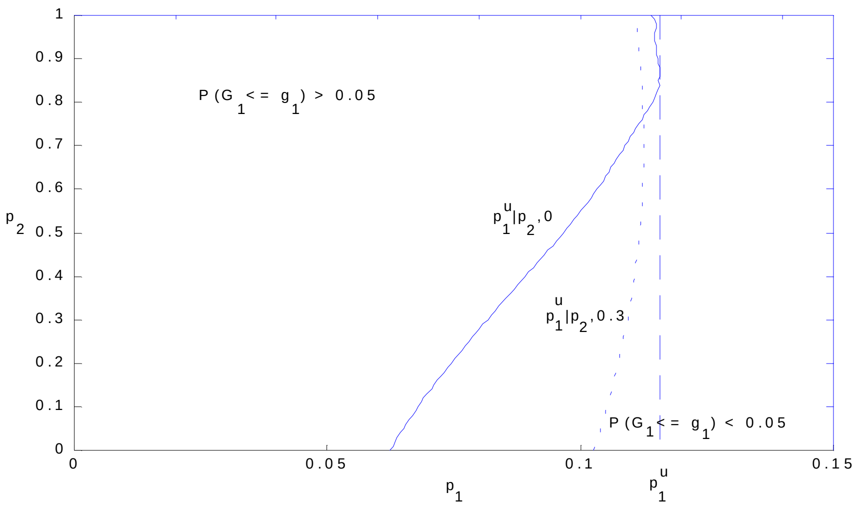
Figure 2. Graph of $p_1^u j p_2; p_4$ for $g_1 = 0{:}042$



Subsequently, the maximum of all $p_1^u j p_2, p_4$ over all possible values of p_2 and p_4 has to be determined. This maximum was found to be 0.116; it was realized for $(p_2; p_4) = (0{:}862; 0{:}000)$ and -because of the symmetry- for $(p_2; p_4) = (0{:}139; 1{:}000)$. Note that the

p_4 values 0 and 1 are inconsistent with the sample results in table 3 $w = 49$ and $z_1 = 1$; however, this is irrelevant since we are now only interested in the final g_1 value 0.042. Figure 3 shows $p_1^u|p_2,p_4$ for $p_4 = 0$ and the accompanying maximum; $p_1^u|p_2; 0.3$ is shown as well for comparison.

Figure 3. Graph of $p_1^u|p_2; p_4$ for $g_1 = 0.042$; $p_4 = 0$ and $p_4 = 0.3$



It is interesting to compare these results with the numerical findings in MOORS et al. (2000). Their data are obtained by changing the value of z_1 in Table 3 to 0 (and x_1 to 2). Then (2) leads to the ML estimates

$$g_1 = 0.051; \quad g_2 = 0.371$$

Furthermore the upper 95%-confidence limit $p_1^u = 0.120$ was calculated.

In the situation with two error types, similar to the situation with one error type, the upper limit is realized for a very high value of p_2 or p_4. In reality, such high values will not often occur, so the next two sections discuss the Bayesian approach; both are based on RAATS (1999).

3 Bayesian approach for one error type

In the one error type situation, the model contains two unknown parameters: p_1 (the error rate in the population), and p_2 (the probability that the auditor misses an error). In the Bayesian approach these two unknown parameters p_1 and p_2 are viewed as realizations of random variables P_1 and P_2. Their prior distribution represents the researcher's knowledge before the sample results are obtained. A logical choice for the marginal prior distributions of P_1 and P_2 is the beta distribution, as the conjugated distribution of the binomial sample results. Further, independence of P_1 and P_2 seems reasonable (the quality of the population is independent of the quality of the auditor), so that the joint prior distribution of P_1 and P_2 is the product of two beta distributions:

$$L(P_1; P_2) \propto p_1^{\alpha-1}(1-p_1)^{\beta-1}p_2^{\gamma-1}(1-p_2)^{\delta-1} \tag{14}$$

The prior knowledge about p_1 (p_2) is reflected by the parameters α and β (γ and δ).

In combination with the binomial sample results (1) this leads to the following posterior distribution of (P_1,P_2):

$$\begin{aligned} L(P_1; P_2|y; y_1; x_2) \propto & \\ \sum_{k=0}^{n-m-x_2} & \Big[(-1)^k \binom{n-m-x_2}{k} p_1^{y+x_2+\alpha+k-1}(1-p_1)^{m-y+\beta-1} \cdot \\ & p_2^{y_1+\gamma-1}(1-p_2)^{y-y_1+x_2+\delta+k-1}\Big] \end{aligned}$$

Integrating over P_2 gives the marginal posterior distribution of the main parameter P_1:

$$\begin{aligned} L(P_1|y; y_1; x_2) \propto & \\ \sum_{k=0}^{n-m-x_2} & \Big[(-1)^k \binom{n-m-x_2}{k} p_1^{y+x_2+\alpha+k-1}(1-p_1)^{m-y+\beta-1} \cdot \\ & B(y_1+\gamma; y-y_1+x_2+\delta+k)\Big] \end{aligned} \tag{15}$$

with $B(a; b) = \int_0^1 x^{a-1}(1-x)^{b-1}dx$

As point estimate $\hat{p}_1$ for p_1 in the Bayesian approach we take the mode of the marginal posterior distribution of P_1; its 0.95-quantile is the Bayesian upper 95% confidence limit $\hat{p}_1^u$. Note that by integrating over P_2, all different values of p_2 are taken into consideration, and not only the worst values as in the classical approach. Hence, $\hat{p}_1^u$ will in general be lower than p_1^u.

An important feature of the Bayesian approach is the choice of the prior distribution parameters. If no specific prior knowledge is available, all possible values of (p_1, p_2) can be considered as equally probable; this leads to the so so-called non-informative prior, defined by $\alpha = \beta = \gamma = \delta = 1$. The choice $\beta > \alpha$ e.g. reflects the researcher's belief that lower values of P_1 are more likely. For simplicity, $\alpha = \gamma = 1$ will be chosen throughout; for β and δ the values 1 and 5 will be considered.

The Bayesian approach is now applied to the practical example of one error type, mentioned in Section 2. For the observed data

$$n = 500; \quad m = 53; \quad y = 3; \quad y_1 = 1; \quad x_2 = 14$$

and the non-informative prior, the posterior (15) becomes

$$L(P_1|y = 3; y_1 = 1; x_2 = 14) \propto (1 - p_1)^{50} \left[\sum_{k=0}^{433} (-1)^k \binom{433}{k} p_1^{17+k} B(2; 17 + k) \right]$$

Figure 4 shows this distribution; the Bayes estimates b_1 an b_1^u are indicated.

Figure 4. Marginal posterior distribution P_1

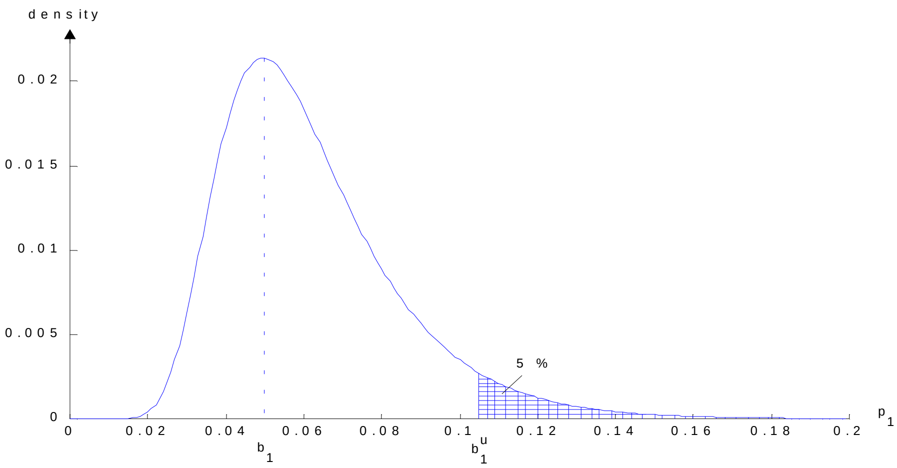


Table 4 summarizes these Bayes estimates for four different priors; for comparison, the classical estimates, mentioned in Section 3 are added.

Table 4. Point estimates and upper limits for p_1; $\alpha = \delta = 1$

Parameters prior		Bayes estimates	
β	γ	b_1	b_1^u
1	1	.050	.105
1	5	.048	.101
5	1	.042	.075
5	5	.042	.073
Classical estimates		.051	.120

All Bayesian estimates are lower than the corresponding classical results. For the upper limits, this is caused by the additional information represented in the prior. Especially prior knowledge about the quality of the auditor has a large impact on the estimates; the researcher's belief that p_2 is low leads to a considerable reduction of b_1 and b_1^u. Prior knowledge about p_2 has a larger influence on the results than prior knowledge about p_1, because there is less sample information concerning p_2.

4 Bayesian approach for two error types

The model with two error types contains, besides the two unknown parameters (p_1 and p_2) of the one error type situation, a third unknown parameter p_4 (the probability that the auditor makes up an error). Independence of P_1 and (P_2, P_4) seems reasonable (the quality of the population is independent of the quality of the auditor), but independence of P_2 and P_4 is questionable. Nevertheless, this assumption is made to simplify the calculations. Starting from marginal beta distributions, the joint prior distribution of P_1, P_2 and P_4 then reads:

$$L(P_1, P_2, P_4) \propto p_1^{\alpha-1}(1-p_1)^{\beta-1} p_2^{\delta-1}(1-p_2)^{\epsilon-1} p_4^{\eta-1}(1-p_4)^{\theta-1} \tag{16}$$

In combination with the binomial sample results (4), this leads to the following joint posterior distribution:

$$\begin{aligned} &L(P_1, P_2, P_4 | y, y_1, z_1, x_2) \propto \\ &\quad p_2^{y_1+\delta-1}(1-p_4)^{m-y-z_1+\theta-1} \sum_{j=0}^{n-m-x_2} \sum_{k=0}^{x_2+j} \Big[(-1)^j \binom{n-m-x_2}{j} \binom{x_2+j}{k} \Big(\\ &\quad p_1^{y+k+\alpha-1}(1-p_1)^{m-y+x_2+j-k+\beta-1}(1-p_2)^{y-y_1+k+\epsilon-1} p_4^{z_1+x_2+j-k+\eta-1} \end{aligned}$$

Integrating over the nuisance variables P_2 and P_4 leads to the marginal posterior distribution of the main parameter P_1 :

$$\begin{aligned} &L(P_1 | y, y_1, z_1, x_2) \propto \\ &\sum_{j=0}^{n-m-x_2} \sum_{k=0}^{x_2+j} \Big[(-1)^j \binom{n-m-x_2}{j} \binom{x_2+j}{k} p_1^{y+k+\alpha-1}(1-p_1)^{m-y+x_2+j-k+\beta-1} \Big] \\ &\quad B(y_1+\delta, y-y_1+k+\epsilon) B(z_1+x_2+j-k+\eta, m-y-z_1+\theta)] \end{aligned} \tag{17}$$

As point estimate $\hat{p}_1$ for p_1 and upper 95% confidence limit $\hat{p}_1^u$, we take again the mode and the 0.95-quantile of the marginal posterior distribution of P_1.

The Bayesian approach is applied to the example of Section 2.3; some additional numerical results are given in Section 5. Using the non-informative prior in combination with the sample results in Table 3, (17) can be simplified to:

$$\begin{aligned} &L(P_1 | y=3, y_1=1, z_1=1, x_2=14) \propto \\ &\sum_{j=0}^{433} \sum_{k=0}^{14+j} (-1)^j \binom{433}{j} \binom{14+j}{k} p_1^{3+k}(1-p_1)^{64+j-k} B(2, 3+k) B(16+j-k, 50) \end{aligned}$$

Figure 5 shows the marginal posterior distribution and the Bayes estimates $\hat{p}_1$ and $\hat{p}_1^u$.

Figure 5. Marginal posterior distribution P_1

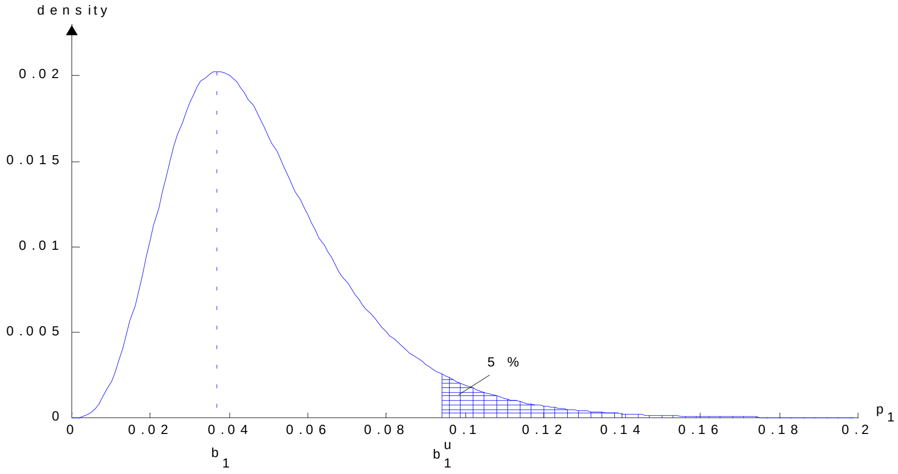


Table 5 contains the classical results calculated in section 2 and the Bayesian results for eight di¤erent priors.

Table 5. Point estimates and upper limits for p_1; ® = ± = ´ = 1

Parameters prior			Bayes estimates	
—	2	µ	b_1	b_1^u
1	1	1	.037	.094
1	1	5	.037	.094
1	5	1	.031	.063
1	5	5	.031	.063
5	1	1	.036	.089
5	1	5	.036	.089
5	5	1	.030	.061
5	5	5	.030	.061
Classical estimates			.042	.116

As in the situation with one error type, all Bayesian estimates are lower than the corresponding classical results and again prior knowledge about p_2 has a larger impact on the results than prior knowledge about p_1. Prior knowledge about p_4 hardly has any impact although this parameter, just like p_2; concerns the quality of the auditor. The explanation is that there is much more sample information on p_4: this parameter is estimated from the $w + z_1 = 50$ correct records in the double-checked sample, and p_2 from only the $y = 3$ incorrect values.

5 Conclusions

In the previous sections both the classical approach and the Bayesian approach of two models for the repeated audit control have been discussed. The calculations were illustrated for one numerical situation. Table 6 shows some more results, for slightly different sample outcomes; appendix B presents some results for small sample sizes.

Table 6. Results of the different approaches

Model	n	m	x	x_2	y	y_1	z_1	Classical g_1	Classical p_1^u	Bayesian b_1	Bayesian b_1^u
Flawless sample check	500	-	16	-	-	-	-	.032	.048	.035	.048
Repeated audit control with one error type	500	53	-	14	3	1	-	.051	.120	.050	.105
	500	53	-	14	2	0	-	.032	.092	.038	.077
Repeated audit control with two error types	500	53	-	14	3	1	0	.051	.121	.043	.099
	500	53	-	14	3	1	1	.042	.116	.037	.094

The most striking feature of this table is the increase of the upper limits for all double check models; even if the expert finds not a single additional error (line 3) p_1^u and b_1^u are 90 and 60%, respectively, larger than when the auditor is assumed to be infallible (line 1).

Lines 2 an 4 represents the empirical data found in Dutch social security payments, where the first auditor made up no errors, but missed one error. In line 2 the model includes only the possibility of missing errors, in line 4 the possibility of making up errors is considered as well. Extending the model with this second error type has not much influence on the classical results, while the Bayesian estimates decrease.

Note that the Bayesian upper limits are generally smaller than the classical ones, but exceptions occur. This can be explained as follows, for example for the one error type situation. Introduce the Bayesian upper limit $b_1^u|p_2$ for a given value of p_2: Then $b_1^u|p_2 < p_1^u|p_2$ will hold, unless the prior distribution of p_1 is concentrated around (much) higher values than the sample information. Now, b_1^u is obtained by averaging $b_1^u|p_2$ with respect to the posterior distribution of p_1; while $p_1^u = \max_{p_2}(p_1^u|p_2)$ considers the worst case. Consequently, only exceptionally b_1^u will exceed p_1^u, that means, for the cases considered here, in particular for the non-informative prior.

The models discussed in this paper consider rather elementary situations, that deviate from practical auditing conditions in two main respects. First of all,

- In practice, the total size of all errors will be of even greater importance than the error rate p_1: hence the size of individual errors ('taintings' in auditors' parlance) will have to be taken into account. This introduces a continuous analysis instead of the right/wrong approach.

2 The previous research started from random sampling. However, in auditors' practices, selection with probabilities proportional to the recorded values ('monetary unit sampling' or MUS) is applied frequently. Hence this sampling method will have to be investigated as well.

In the Bayesian approach it was assumed that the probability of missing an error is independent of the probability of making up an error. However, this assumption is questionable and it would be interesting to repeat the above investigations without assuming independence.

Finally, a number of more theoretical issues remain. For example, according to LEHMANN (1959, p.176) no uniformly most accurate confidence set will in general exist in the presence of nuisance parameters, as in our case, but perhaps our method of constructing upper limits can be improved.

Acknowledgments

This paper is based on the first author's master thesis, written during her stay at Enterprise Risk Services, Deloitte & Touche; this thesis won an honourable mention in the 2000 VVS Master thesis contest. We are grateful for the support and comments of Paul C. van Batenburg and Anita Fokker. Very useful remarks from Ben van der Genugten (Tilburg University) are gratefully acknowledged.

A Symmetry of the ML-estimator

In the case of two possible error types, it will be shown here by means of three consecutive lemmas that the distribution of the ML estimator G_1 for p_1 is symmetric with respect to $(p_2; p_4) = (0.5; 0.5)$, that is: $L(G_1|p_1; p_2; p_4) = L(G_1|p_1; 1 - p_2; 1 - p_4)$.

Introduce $V = (Y; Y_1; Z_1; X_2)$; define the functions $f : R^4 \to R^4$ and $h : [0; 1]^3 \to [0; 1]^3$ by

$$f(v) = f(y; y_1; z_1; x_2) = (y; y - y_1; m - y - z_1; n - m - x_2) \text{ and}$$

$$h(v) = h(p_1; p_2; p_4) = (p_1; 1 - p_2; 1 - p_4)$$

and define the set A_c for all $c \in [0; 1]$ by

$$A_c = \{v : g_1(v) = c\}$$

Note that $f \circ f$ and $h \circ h$ are both the identity.

Lemma 1 $f(A_c) = A_c$

Proof: The special case $v = (y; y; 0; x_2)$ implies $f(v) = (y; 0; m - y; n - m - x_2)$ and $g_1(v) = g_1(f(v)) = \dfrac{y}{m}$

In the general case, $g_1(v) = g_1(f(v))$ can be proved similarly. Hence $v \in A_c$ implies $f(v) \in A_c$, and vice versa.

Lemma 2 $Pr(V = v|p) = Pr(V = f(v)|h(p))$

Proof: By direct verification, using (4).

Lemma 3 $Pr(G_1 = c|p) = Pr(G_1 = c|h(p))$

Proof:

$$\begin{aligned} Pr(G_1 = c|h(p)) &= Pr(V \in A_c|h(p)) \\ &= Pr(V \in f(A_c)|h(p)) \\ &= Pr(V \in f(A_c)|p) \\ &= Pr(G_1 = c|p) \end{aligned}$$

where the second equality follows from Lemma 1 and the third from Lemma 2.

B Estimates and con…dence limits for p_1 (n = 50)

Table A. Flawless sample check

Sample results		Classical		Bayesian			
				no prior knowledge		prior knowledge	
n	x_2	g_1	p_1^u	b_1	b_1^u	b_1	b_1^u
50	4	.080	.174	.080	.171	.074	.159
50	5	.100	.199	.115	.195	.093	.182
50	6	.120	.223	.135	.219	.111	.204

Table B. Repeated audit control with one error type

Sample results					Classical		Bayesian			
							no prior knowledge		prior knowledge	
n	m	x_2	y_1	y	g_1	p_1^u	b_1	b_1^u	b_1	b_1^u
50	20	2	0	2	.080	.222	.093	.213	.087	.189
50	20	2	1	3	.131	.289	.132	.278	.117	.237
50	20	2	0	1	.060	.216	.071	.186	.065	.161
50	20	2	1	2	.106	.283	.109	.250	.094	.208
50	20	2	0	0	.040	.160	.049	.157	.044	.132
50	20	2	1	1	.088	.226	.085	.221	.071	.178
50	20	3	0	3	.120	.283	.136	.262	.129	.240
50	20	3	1	4	.172	.344	.176	.325	.161	.289
50	20	3	0	2	.100	.283	.114	.236	.108	.214
50	20	3	1	3	.150	.344	.153	.298	.138	.261
50	20	3	0	1	.080	.222	.092	.210	.086	.188
50	20	3	1	2	.128	.289	.130	.271	.116	.234
50	20	3	0	0	.060	.216	.070	.182	.065	.160
50	20	3	1	1	.107	.283	.107	.243	.093	.206

Table C. Repeated audit control with two error types

Sample results						Classical		Bayesian			
								no priorknowledge		prior knowledge	
n	m	x_2	z_1	y_1	y	g_1	p_1^u	b_1	b_1^u	b_1	b_1^u
50	20	2	0	0	2	.080	.228	.081	.204	.075	.179
50	20	2	0	1	3	.131	.291	.122	.217	.107	.229
50	20	2	1	0	1	.040	.164	.043	.163	.040	.139
50	20	2	1	1	2	.091	.238	.085	.234	.073	.191
50	20	2	2	0	0	.000	.139	.000	.114	.000	.091
50	20	2	2	1	1	.051	.216	.046	.193	.038	.148
50	20	3	0	0	2	.100	.283	.096	.222	.091	.200
50	20	3	0	1	3	.150	.344	.137	.287	.124	.250
50	20	3	1	0	1	.050	.216	.051	.176	.049	.156
50	20	3	1	1	2	.100	.283	.094	.244	.085	.209
50	20	3	2	0	0	.000	.139	.000	.121	.000	.103
50	20	3	2	1	1	.050	.216	.049	.197	.044	.162
50	20	3	0	0	3	.120	.286	.122	.252	.116	.230
50	20	3	0	1	4	.178	.347	.164	.318	.150	.280
50	20	3	1	0	2	.080	.228	.085	.213	.080	.191
50	20	3	1	1	3	.132	.295	.128	.281	.115	.243
50	20	3	2	0	1	.040	.164	.044	.170	.041	.148
50	20	3	2	1	2	.092	.239	.089	.241	.078	.202
50	20	3	3	0	0	.000	.169	.000	.118	.000	.097
50	20	3	3	1	1	.052	.216	.047	.197	.040	.160

References

COX, D.R. and D.V. HINKLEY (1974), Theoretical statistics, Chapman and Hall, London.

LEHMANN, E.L. (1959), Testing statistical hypotheses, Wiley, New York.

MOORS, J.J.A. (1999). Double checking for two error types, CentER Discussion Paper 9923, KUB.

MOORS, J.J.A., B.B. van der Genugten and L.W.G. Strijbosch (2000). Repeated audit controls, Statistica Neerlandica 54, p.p. 3-13

RAATS, V.M. (1999). Herhaalde steekproefcontrole, Master thesis Tilburg University